

Claims:

1. A method for managing a web site registered in a search engine, said method comprising the steps of:

5 receiving web site information of the registered web site, classifying the web site information by predetermined fields, and recording the classified web site information in a database;

reading a source file constituting a web page of the registered web site;

analyzing the read source file;

10 determining, based on a predetermined basis, whether or not the registered web site is a deceptive site; and

performing a control operation to perform predetermined processing on the registered web site if the web site is determined to be a deceptive site.

2. The method according to claim 1, wherein the source file is an HTML (Hypertext Markup Language) document.

15 3. The method according to claim 2, wherein the predetermined basis is whether or not the HTML document includes a character string of the same color as background color of the web page.

4. The method according to claim 2, wherein the predetermined basis is whether or not a redirection tag in the HTML document includes a character string.

20 5. The method according to claim 2, wherein the predetermined basis is whether or not the length of a title tag included in the HTML document is more than a predetermined numerical value or whether or not the number of title tags included therein is more than one.

25 6. The method according to claim 2, wherein the predetermined basis is whether or not the length of a character string in a meta tag included in the HTML document is more than a predetermined numerical value.

7. The method according to claim 2, wherein the predetermined basis is whether or not a character string exists in a frame tag in the HTML document.
8. The method according to claim 2, wherein the predetermined basis is whether or not the length of a character string included in a form tag in the HTML document is 5 more than a predetermined numerical value.
9. The method according to claim 2, wherein the predetermined basis is whether or not the length of the same character strings in a div tag in the HTML document is more than a predetermined numerical value.
10. The method according to claim 2, wherein the predetermined basis is whether 10 or not an href tag in the HTML document includes a character string other than a URL (Universal Resource Locator).
11. The method according to claim 2, wherein the predetermined basis is whether or not the HTML document includes links which link web pages in the same web site, the number of said links being more than a predetermined number.
- 15 12. The method according to claim 2, wherein the predetermined basis is whether or not the HTML document includes a character string whose font size is zero.
13. The method according to claim 2, wherein the predetermined basis is whether or not the length of a character string included in an img tag in the HTML document is more than a predetermined numerical value.
- 20 14. The method according to claim 2, wherein the predetermined basis includes at least two of the predetermined bases defined in claims 3 to 13.
15. The method according to claim 14, said step of determining whether or not the registered web site is a deceptive site includes the steps of:
maintaining predetermined weighted values corresponding respectively to the

predetermined bases;

calculating respective point values of the predetermined bases according to predetermined point calculation methods corresponding respectively to the predetermined bases;

5 calculating respective intermediate values of the predetermined bases by multiplying the calculated point values respectively by the weighted values corresponding to the predetermined bases corresponding to the calculated point values;

10 obtaining a sum of the calculated intermediate values of the predetermined bases; and

determining whether or not the sum of the respective intermediate values is more than a predetermined value, and determining that the web site is a deceptive site if the sum is more than the predetermined value.

15 16. The method according to claim 2, wherein the predetermined basis is whether or not a combined color value of RGB components of pixels included in an image file contained in an img tag in the HTML document is more than a predetermined value.

17. The method according to claim 1, wherein the database includes a web site registrant field, and

20 wherein said step of performing the control operation to perform the predetermined processing comprises the steps of:

obtaining information of a registrant of the web site by searching the web site registrant field of the database;

extracting contact information of the registrant from the registrant information of the web site; and

25 transmitting a message to a location corresponding to the extracted contact information.

18. The method according to claim 17, wherein the contact information is a mobile terminal number or an email address of the registrant of the web site, and

wherein said step of transmitting the message comprises the step of

controlling an email server to send an email to the email address or the step of controlling an SMS server to send an SMS message to the mobile terminal number.

19. The method according to claim 1, wherein the database includes a web site registrant field, and

5 wherein said step of performing the control operation to perform the predetermined processing comprises the steps of:

obtaining information of a registrant of the web site by searching the web site registrant field of the database;

10 extracting a URL of a different web site registered by the registrant from the registrant information of the web site; and

reading a source file constituting a web page of a web site connected through the URL;

analyzing the read source file;

15 determining, based on a predetermined basis, whether or not the web site is a deceptive site; and

performing a control operation to perform predetermined processing on the registered web site if the web site is determined to be a deceptive site.

20. A computer-readable recording medium in which a program for performing the method defined in any one of claims 1 to 19 is recorded.

20 21. A system for managing a web site registered in a search engine, the system comprising:

an interface module for performing data communication with at least one terminal;

25 a web site registration module for receiving a web site registration request including web site information of a predetermined web site from said at least one terminal and classifying the web site information by predetermined fields;

a database for classifying and storing a predetermined keyword corresponding to the web site and the web site information;

a web site analysis module for extracting a source file constituting a web page

of the web site, and analyzing the extracted source file; and
a web site management module for determining, based on a predetermined
basis, whether or not the web site is a deceptive site.